融合元数据及 Attention 机制的深度联合学习推荐 *

张全贵，李志强，张新新，曹志强

(辽宁工程技术大学 电子与信息工程学院, 辽宁 葫芦岛 125105)

摘　要：融合元数据的协同过滤推荐即混合推荐算法是目前推荐系统领域研究的热点，能一定程度地解决数据稀疏及冷启动等问题。但融合元数据现有的建模方法大多数建立于用户/项目属性权重相同的情景下，以致于用户项目间重点关系表达不显著，难以获得较好的推荐性能。针对上述问题，提出一种融合元数据及 Attention 机制的深度联合学习推荐方法。它利用双深度网络联合学习，其中一个网络基于隐反馈数据实现矩阵非线性分解以学习用户/项目个性化关系，另一个利用 Attention 机制自动捕捉用户/项目关键属性对推荐工作的影响，通过赋予不同属性权重凸显的用户偏好关系建模辅以扩展模型。实验结果表明，所提出的推荐算法在 MovieLens 100K 和 MovieLens 1M 两个公开数据集上均表现出较为优越的推荐性能。

关键词：元数据；属性权重；Attention 机制；深度联合学习；非线性分解

中图分类号：TP182　　doi: 10.3969/j.issn.1001-3695.2018.04.0285

Deep Joint learning recommendation based on metadata and Attention mechanism

Zhang Quangui, Li Zhiqiang, Zhang Xinxin, Cao Zhiqiang

(*School Electronics & Information Engineering Liaoning Technical University*, *Huludao Liaoning* 125105, *China*)

Abstract: Collaborative filtering recommendation, which combines metadata, is a hot topic in recommender systems, which can solve data sparsity and cold-start problems to some extent. However, most of the existing modeling methods of combines metadata are based on the same user / item attribute weights, so that the key relationships between users and items are not significant, and it is difficult to obtain better recommendation performance. To solve the above problems, this paper proposes a method of deep joint learning recommendation based on metadata and attention mechanism. It uses double deep network learning, one of the networks implements matrix nonlinear decomposition based on implicit feedback data to learn user/project personalization relationships. , and another network automatically capture affect the user/item key attributes to recommend by using Attention mechanism, through the user preference relation with weights of different attributes modeling highlights the extended model. Experimental results show that the proposed recommendation algorithm has better recommendation performance on two public datasets of MovieLens 100K and MovieLens 1M.

Key words: metadata; attribute weight; Attention mechanism; deep joint learning; nonlinear factorization

0　引言

推荐系统是通过估测用户需求来缓解信息超载问题的一种技术，被广泛应用在新闻网站、社交媒体及电子商务等领域。美国专家 Resnick 于 1997 率先提出“推荐系统”这一概念，之后推荐系统在学术界和工业界均引起巨大的关注，相继不同的推荐机制被研究者提出，并成为推荐系统学术领域的研究热点。

在现存方法中，协同过滤推荐已经被广泛研究和扩展，如 Koren、Paterek、Lee 等人[1~3]提出的基于传统的矩阵分解模型的协同过滤算法。此后研究人员还试图通过元数据建模以扩展矩阵分解模型即所谓的混合推荐，能在一定程度上解决稀疏性和冷启动问题。如 Fu 等人[4]从不同视角度量元数据中用户/项目潜在的个性化关系，利用这种关系扩展矩阵分解模型以实现推荐；Li 等人[5]提出物品耦合关系的相似性计算方法，然后将其整合到矩阵分解方法以提升推荐性能。但以上已有的方法均属于浅层模型，不能发现更为抽象的复杂交互关系。

而如今利用深度学习技术建立深层模型已广泛地应用在推荐系统领域[6]，它可以有效地捕获非线性和复杂的用户/项目关系以表示更复杂、抽象的数据结构。如 Chang 等人[7]人通过在传统的协同过滤算法中应用神经网络算法达到了提高推荐质量

收稿日期：2018-04-19；**修回日期：**2018-06-15　　**基金项目：**国家留学基金资助项目（留金法［2015］5104）

作者简介：张全贵（1978-），男，河北秦皇岛人，副教授，博士，主要研究方向为推荐系统、机器学习、数据挖掘等（zhqgui@126.com）；李志强，男，辽宁锦州人，主要研究方向为推荐系统，机器学习等；张新新（1994-），女，辽宁朝阳人，硕士，主要研究方向为机器学习、推荐系统、数据挖掘等；曹志强（1988-），男，河南驻马店市人，硕士，主要研究方向为图像处理与模式识别.

的目的。文献中的 CCCFNet 也是双网络（协同过滤和内容信息）[8]，它模拟用户项目之间的相互作用在顶层的点积实现推荐工作。Guo 等人[9]是一个无缝集成的分解机和 MLP 的端到端的模型。它通过因子分解机利用深层神经网络和低阶相互作用对高阶特征相互作用进行建模。He 等[10]也利用多层深度网络替代矩阵线性分解机制实现非线性分解，一定程度地解决了线性 MF 的局限性。但上述几种模型捕捉的是全部属性间的交互，但在实际应用中，部分交互是没有必要的。例如图 1 所示，学习用户/项目群体划分时，用户职业、项目类别等是关键的划分特征，而性别、地域等信息是非重要划分成分。另一个方面，以上现有的融合元数据建模方法大多数建立于用户/项目属性权重相同的情景下，即属性的值从模型的开始训练到结束都是不变的，忽视了用户/项目不同属性对推荐工作的影响，以致于用户项目间重点关系表达不显著，大大地降低了推荐模型的性能。因此研究人员逐渐认清无论模型是如何构建的，属性都在其中起着决定性的作用。

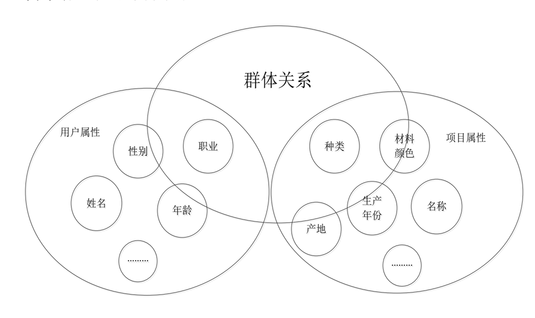


图 1　不同属性占比影响

近些年，Attention 网络被研发并逐渐成为捕获全局信息的技术手段。因此，基于 Attention 机制的神经网络在各个领域成为一个热点，是解决上述问题的有效方法之一。如 Luong 等[11]在自然语言处理领域使用 Attention 机制对所有词分配不同权值实现翻译工作。google mind 团队[12]在计算机视觉领域使用 Attention 机制自动捕捉图像局部特征实现图像分类。文献[13]提出一种基于注意力机制的融合性深度神经网络，将 Twitter 和用户兴趣之间的用户、作者、用户兴趣相似性信息融合作为学习属性实现预测用户的转发行为。但基于 attention 机制的应用在推荐领域中却少之又少。接下来，我们将这种思想应用于推荐系统领域中，即假设用户具有五个属性值 e1、e2、e3、e4、e5，它们对推荐作用的权重为 w1、w2、w3、w4、w5，在不加入 attention 情况下，权重 w1~w5 是相同的，则总体的影响作用为 e1*w1+e2*w2+e3*w3+e4*w4+e5*w5。反之加入时，attention 模型根据数据学习出当前用户的属性的对当前推荐的不同影响因子 m1、m2、m3、m4、m5，最后总体的影响作用 e1*m1*w1+e2*m2*w2+e3*m3*w3+e4*m4*w4+e5*m5*w5。

据此，本文提出融合元数据及 Attention 机制的深度联合学习推荐方法，将新型 Attention 网络特征学习推理和深度网络分解机功能相结合，形式化用于协同过滤的建模方法来解决上述研究问题。我们通过利用双深度网络联合学习实现预测推荐，其中一个网络基于隐式反馈数据实现矩阵非线性分解以学习用户/项目个性化关系，另一个利用 Attention 机制自动捕捉用户/项目关键属性对推荐工作的影响，通过赋予不同属性动态权重凸显用户偏好关系，能有效地挖掘出属性有用的隐含特征，从而减少特征冗余、盲目交互等带来的消极影响。本文算法在 MovieLens 100K 和 1M 两个公开数据集上进行验证。最终实验结果表明所提出的算法均表现出较为优越的推荐性能。

1　算法概述

融合元数据及 Attention 机制的深度联合学习推荐算法（简称 DJRMA）是将 Attention 网络特征学习推理和深度网络分解机相结合，通过学习个性化关系及挖掘元数据隐含关系来提升推荐模型的性能。如图 2 所示，左侧分解机制网络模型基于隐反馈数据且利用深度网络代替矩阵线性分解实现非线性分解，以学习用户/项目的个性化偏好。右侧模型利用 Attention 机制自动捕捉用户/项目关键属性对推荐工作的影响，通过赋予不同属性权重凸显的用户偏好关系建模辅助扩展模型。其核心思想是在原有的用户/项目交互矩阵基础上加上已排除干扰噪声的用户个性化向量与项目个性化向量乘积重构后的用户/项目交互矩阵，进而形成新的联合交互矩阵以实现预测推荐。

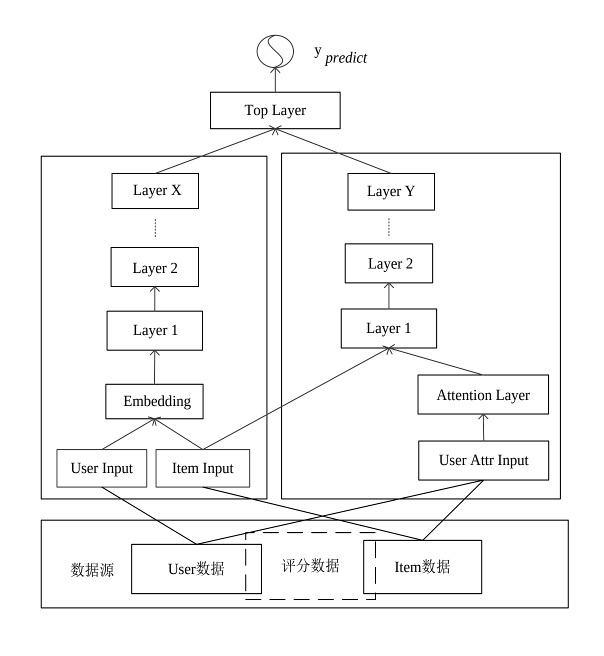


图 2　本文推荐算法实验模型

1.1　深度分解机制

深度分解机制是基于隐式反馈数据利用多层深度网络替代矩阵线性分解机制实现非线性分解，以学习用户项目的个体个性化关系。底部输入层分别由描述用户 u 和项目 i 的两个特征向量组成，它们可被定制以支持用户和项目的广泛建模。在本文中，为了更方便说明问题，我们只采用用户/项目 ID 作为输入通用特征，而模型根据实际可以有不同的输入。其次，输入层的上方是嵌入层，如图 3 所示其将用户/项目 ID 这种稀疏向量映射成稠密向量，获得到的用户/项目特征向量可以被看作在潜在因素模型上下文中用户/项目潜在的特征。

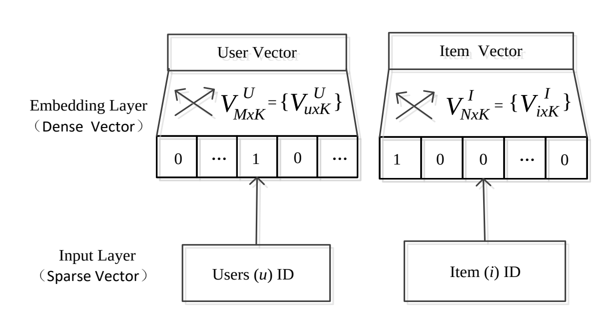


图 3　稀疏向量映射稠密向量

因此，定义深度分解模型如下：

$$\hat{y}^1_{ui}=f\left(V^U_u,V^I_i\middle|U,I,\theta\right) \tag{1}$$

其中：$U\in R^{M\times K}$ 和 $I\in R^{N\times K}$ 分别表示用户和项目的潜在因素矩阵，M、N 和 K 为用户、项目和潜在特征数量；θ 表示 f 函数的模型参数。接下来，稀疏向量用户（项目）ID 通过嵌入层映射成稠密向量，即用户/项目的特征向量。我们使用 V^U_u 和 V^I_i 分别为用户 u 和项目 i 的特征向量。其前向过程为

$$z_1=\psi_1\left(V^U_u,V^I_i\right)=\begin{bmatrix}V^U_u\\V^I_i\end{bmatrix}, \tag{2}$$

$$\psi_2(z_1)=\sigma\left(W_2'^T z_1+b_2'\right), \tag{3}$$

……

$$\psi_H(z_{H-1})=\sigma\left(W_H'^T z_{H-1}+b_H'\right) \tag{4}$$

其中：H 是层的深度，σ 是激活函数。z_H、W_H' 和 b_H' 是第 H 层的输出、模型权重和偏置项。经过 L 层网络学习后，最终左侧模型的映射函数为

$$\begin{aligned}\hat{y}^1_{ui}&=\sigma(\psi_L(z_{L-1}))\\&=\sigma\left(W_L'^T\left(\sigma\left(\ldots\ \sigma\left(W_2'^T\begin{bmatrix}V^U_u\\V^I_i\end{bmatrix}+b_2'\right)\ldots\right)\right)+b_L'\right)\end{aligned} \tag{5}$$

1.2　基于 Attention 特征学习推理机制

Attention model（注意力模型）是一种模拟人脑注意力的模型，即在特定的时刻人脑对事物的注意力会集中在某一特定的地方，忽略其他部分的特点，是一种人脑资源分配的模型。它是对重要部分分配较多的注意力，对于其他部分分配较少的注意力，合理利用人脑的计算资源，并且还可以去除非关键因素对人脑的影响。本文引入这种思想，利用 Attention 机制自动捕捉用户/项目关键属性对推荐工作的影响，通过赋予不同属性权重来凸显用户偏好关系，然后通过多层深度网络建模来辅助扩展上述非线性分解模型，有效地防止元数据隐含信息的丢失。

如图 4 所示，基于 Attention 特征学习推理机制将用户属性特征表示为 one-hot 编码形式向量，并输入到 Attention 网络中自动捕捉用户/项目关键属性，并推理出对推荐工作的影响度。本文实验中使用电影 ID 作为输入，用户属性则利用性别、年龄及职业信息作为输入。假设使用 V^U_u 用户 u 的特征向量，A_u 表示用户 u 的属性，$V^U_u=\{A_{u1},A_{u2},A_{u3},\ldots A_{ug}\}$，其中 g 表示用户属性数量，M 是用户的数量。

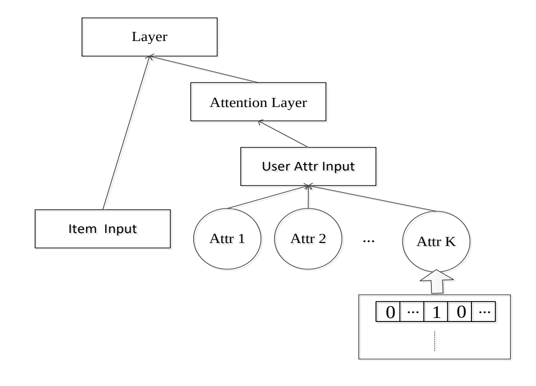


图 4　基于 Attention 局部特征学习推理机制

如图 5 所示，可定义编码隐状态 $(h_1,\ldots,h_{T_A})$ 和解码隐状态 $(d1,\ldots,d_{T_B})=(h_{T_A+1},\ldots,h_{T_A+T_B})$。在输入的用户 g 个属性上，为了计算每个输出时间 t 的注意力权重的属性影响因子矩阵，定义为

$$e_i^t=f(h_i,d_t)=v_a{}^T tanh(W_a h_i+W_b d_t) \tag{6}$$

(7)

其中：f 表示一种模式函数，W_a，W_b，v 是模型的学习参数。向量 e^t 的长度是 T_A（即输入的用户属性数量）以及它的第 i 项包含一个有多少注意力应该放在第 i 个编码隐藏状态 h_i 上。然后使用 softmax 标准化，得到属性 A_{ui} 在 t 时刻的注意力。最后求和形成权重分配向量，即

$$c^t=\sum_{i=1}^{T_A}a_i^t h_i \tag{8}$$

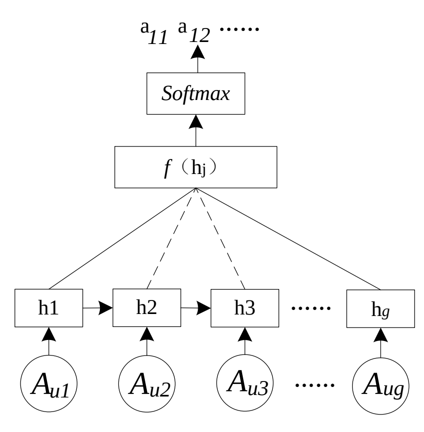


图 5　Attention layer

前向过程为

$$z_1=\psi_1\left(V^U,V^I,c\right)=\begin{bmatrix}cV^U\\V^I\end{bmatrix} \tag{9}$$

$$\psi_2(z_1)=\sigma\left(W_2'^T z_1+b_2'\right) \tag{10}$$

$$\psi_H(z_{H-1})=\sigma\left(W_H'^T z_{H-1}+b_H'\right) \tag{11}$$

其中：H 是层的深度，σ 是激活函数。z_H、W'_H 和 b'_H 是第 H 层的输出、模型权重和偏置项。经过 L 层网络学习后，最终右侧模型的映射函数为

$$\hat{y}_{ui}^2=\sigma(\psi_L(z_{L-1})) \tag{12}$$

模型框架顶层输入来自于两侧模型的输出。由此，我们定义联合预测函数为

$$\hat{y}=\sigma(\alpha\hat{y}_{ui}^1+(1-\alpha)\hat{y}_{ui}^2) \tag{13}$$

其中：α 是决定双深度模型间训练权重的参数，通过深度模型自学习获得。算法将推荐问题形式化一个多分类问题来实现评分预测。采用 Softmax 函数作为最终的激活函数 σ_{out} 以及利用多分类对数损失来学习。通过上面所述设置，我们定义似然函数为

$$P(O,O^-|U,I,\theta)=\prod_{(u,i)\in O}\hat{y}_{ui}\prod_{(u,i)\in O^-}(1-\hat{y}_{ui}) \tag{14}$$

其中：O 表示用户交互集中可观察到的一组交互集合（作为正实例），O^- 为均匀地从未观察到的交互集中取样的一组集合（作为负实例）。将上式取似然负对数作为损失函数可得

$$L=-\sum_{(u,i)\in O\cup O^-}y_{ui}\log\hat{y}_{ui}+(1-y_{ui})\log(1-\hat{y}_{ui}) \tag{15}$$

2　实验及结论分析

2.1　实验数据集

目前，对于推荐领域已经有很多公开的研究。为了验证本文提出的融合元数据及 Attention 机制的深度联合学习推荐算法的性能，将在实验使用 MovieLens 100K 和 MovieLens 1M 两个公开数据集作为实验数据集[14]。MovieLens 数据集由美国明尼苏达大学计算机科学与工程学院 GroupLens 研究组提供。其中每个用户至少对 20 个电影进行评分，数据稀疏程度分别是 0.9369 和 0.9553，并还提供了用户职业、电影流派等辅助信息。MovieLens 100K 包含 943 名用户、1682 部电影、19 种电影流派以及 10 万个 5 星级评价。1M 数据集包含 6040 名用户、3952 部电影、18 种电影流派以及 100 万个 5 星级评价。除此之外，还提供了用户职业、性别等辅助信息。虽然隐式反馈信息要比显式反馈信息更容易获取，但由于目前基于隐式反馈信息并且具有用户元数据信息的公开数据集比较少，因此本文仿照目前常用的方法如文献[15,16]将显式评级数据转换成隐式数据。

2.2　评估方法

本文所提出的算法利用深度网络实现多分类的评分预测工作。而目前推荐领域中被用来度量评分预测工作质量有很多不同的度量指标。为了评价项目推荐的性能，我们采用了 leave-one-out 方法评估，这种 top-K 评价方法已被广泛地应用于文献[17~19]。实验中根据时间戳随机抽取 100 个不与用户进行交互的项目，将测试项目排列在这 100 个项目中。排名列表的性能由命中率（HR）和归一化折扣累积增益（NDCG）[18]来衡量。若无特别说明，我们将这两个指标的排名列表截断 K 均为 10。

命中率 HR 能直观地衡量测试项目是否存在于前 10 项列表中，而 NDCG 是用来衡量排序质量的指标，通过将较高分数指定为顶级排名来计算命中的位置。HR 与 NDCG 定义如下：

命中率（HR）

$$HR@K=\frac{Number\ \ of\ \ Hit@K}{|GT|} \tag{16}$$

归一化折扣累积增益（NDCG）

$$NDCG@K=Z_K\sum_{i=1}^{K}\frac{2^{r_i}-1}{\log_2(i+1)} \tag{17}$$

其中：GT 表示测试项目与算法产生排名列表的集合，Z_K 是正规化，以确保排名列表中存在一个值 1，r_i 是第 i 个项目的评级相关性。在评估过程中我们使用简单的二进制进行表示：如果项目在测试集 $r_i=1$，否则为 0。

2.3　实验环境及参数设置

为了说明本文所提融合元数据及 Attention 机制的深度联合学习推荐模型的优越性，本次实验在 Cluster 一个 node 上实现了所提出的算法（实验环境如表 1 所示）。同时将其与以下具有代表性的经典模型进行实验对比。

a)CBMF[21]:内容加强矩阵分解模型,此模型将元数据信息直接加入到矩阵分解模型中。

b)NeuMF[10]：基于神经网络协同过滤方法，对比辅助信息对推荐效果的影响；

c)ItemKNN[22]：标准的基于项目的协同过滤方法。

表 1　实验环境

软硬件环境	环境配置
操作系统	Linux
主频	3.5GHz
内存	96.00GB
编程语言	Python 3.5
深度学习库	Tensorflow1.2.0、 Keras2.0.5

2.4　实验结果及分析

为了确定参数，本文采用 5 折交叉验证，使用了项目属性电影流派及用户属性性别、年龄及职业等信息，将正例以及对每个正例进行四个负实例的均匀采样作为输入。此后再通过多分类对数损失函数来评估。模型从零开始训练，学习率设置为 0.001，并利用 Adam 经过偏置校正后每一次迭代学习率都有个确定范围以致于使得参数趋于平稳的优点优化模型。网络模型设计遵循塔式网络结构，每个连续较高层的层尺寸减半，层的结构为 128→64→32。中间层激活函数选择 ReLU，因为它支持稀疏激活，使模型不太容易过拟合。首先，我们对本文模型的推荐性能进行了评测，最终实验结果如表 2 所示。

表 1 在 MovieLens 数据集上实验对比（K=10）

Model	DataSet\Metries			
	MovieLens 100K		MovieLens 1M	
	HR	NDCG	HR	NDCG
NeuMF	0.672	0.384	0.705	0.423
ItemKNN	0.605	0.328	0.625	0.362
CBMF	0.632	0.367	0.679	0.393
DJRMA	0.685	0.402	0.712	0.428

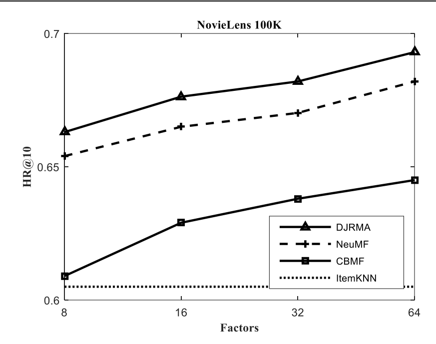


图 6　100K 数据集—命中率

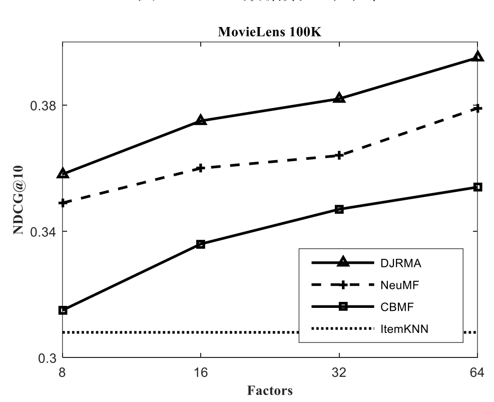


图 7　100K 数据集—归一化折扣累积增益

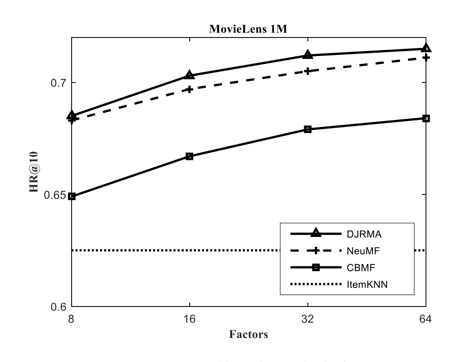


图 8　1M 数据集—命中率

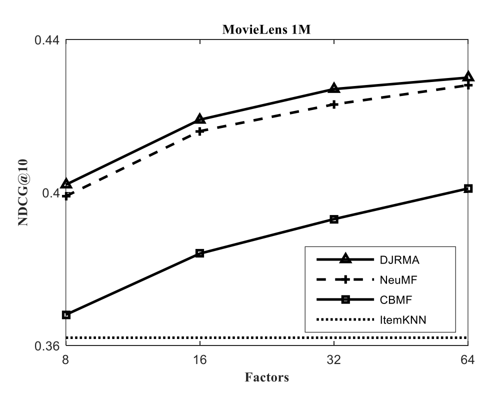


图 9　1M 数据集—归一化折扣累积增益

其次，实验改变层的结构来评测性能，即相当于改变矩阵分解维度。由图 6 与图 7 可知，DJRMA 模型通过增强关键属性影响因子使得推荐性能指标 HR 及 NDCG 进一步增长。在同等设置情况下，100K 数据集 Factors 为 32 时的 HR 相比于 NeuMF、CBMF 和 ItemKNN 分别提升 1.9%、8.3%和 13.2%，NDCG 也提升 4.6%、9.5%和 22.5%。同样通过图 8 和图 9 分析可知，对于在 1M 数据集上，HR 相比于 NeuMF、CBMF 和 ItemKNN 分别提升 0.9%、4.8%和 13.9%，NDCG 也提升 1.1%、8.9%和 18.2%。因此，推荐算法比基线方法表现出较为优越的推荐性能

此外，再次比较图 6 与图 8、图 7 与图 9 可知，DJRMA 推荐模型的推荐指标在 1M 数据集上较优越于 100K。HR 和 NDCG 提升了 3.9%和 6.4%。所以由于训练集增大以及样本的不平衡性，DJRMA 能表现出较强的学习能力，但总体提升率降低，说明样本不平衡性等问题对模型具有一定的影响。

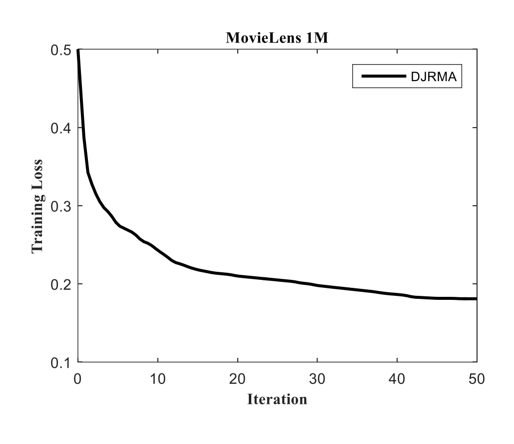


图 10　模型训练损失

图 10 显示在 MovieLens 1M 数据集下不同迭代次数的迭代损失。模型利用隐式反馈实现二进制分类推荐工作，我们通过训练对数损失获取最优的模型参数。由图可知，随着模型迭代次数的增加，模型训练损失依次逐渐减小，直至稳定。实验中最有效的迭代次数约为 10-20 次左右。在后续 MovieLens 100K 实验中，我们也发现其变化趋势与 1M 类似，并且更多的迭代可能使模型饱和，以致于训练损失指标不再提升。

3　结束语

本文将新型 Attention 网络特征学习推理和深度网络分解机功能相结合，提出融合元数据及 Attention 机制的深度联合学习推荐方法，其一实现矩阵非线性分解以学习用户/项目个性化关系，另一个利用 Attention 机制自动捕捉用户/项目关键属性对推荐工作的影响，通过赋予不同属性权重凸显的用户偏好关系建模来扩展模型。该算法通过形式化用于协同过滤的建模方法来解决推荐问题。最终通过双深度网络联合学习实现评分预测。最后，本文算法在 MovieLens 100K 和 1M 两个公开数据集上进行验证。实验结果表明所提出的推荐算法比基线方法不仅仅表现出了较为优越的推荐性能，而且在学习效率上也相差较

小。

在未来工作中，将考虑引入更为丰富的元数据信息，进而挖掘交叉属性间对关系表达的影响因子以改进算法使得推荐质量提高。另外，仍考虑利用残差网络、卷积网络或其它深度模型[23]构建更合理的推荐框架以综合利用不同抽象级别的特征，也是目前推荐领域研究的热点。

参考文献：

[1] Bell R, Koren Y, Volinsky C. The bellkor 2008 solution to the Netflix prize [R]. Korbell Teams Report to Netflix, 2008: 709-714.

[2] Paterek A. Improving regularized singular value decomposition for collaborative filtering [C]// Proc of KDDCup & Workshop, 2007: 319-325.

[3] Lee D D, Seung H S. Learning the parts of objects by non-negative matrix factorization [J]. Nature, 1999, 401: 788-791.

[4] Fu Bin, Xu Guangdong, Cao Longbing, *et al.* Coupling multiple views of relations for recommendation [M]// Advances in Knowledge Discovery and Data Mining. [S. l.] : Springer International Publishing, 2015: 732-743.

[5] Li Fangfang. Non-IID recommender system [D]. Beijing: Beijing Institute of Technology. 2014.

[6] Salakhutdinov R, Mnih A, Hinton G. Restricted Boltzmann machines for collaborative filtering [C]// Proc of International Conference on Machine Learning. New York: ACM Press, 2007: 791-798. ,

[7] Zhang Feng, Chang Huiyou. A collaborative filtering algorithm embedded BP network to ameliorate sparsity issue [C]// Proc of International Conference on Machine Learning and Cybernetics. 2005 1839-1844.

[8] Lian Jianxun, Zhang Fuzheng, Xie Xing, *et al.* CCCFNet: a content-boosted collaborative filtering neural network for cross domain Recommender Systems [C]// Proc of International Conference on World Wide Web Companion. International World Wide Web Conferences Steering Committee, 2017: 817-818.

[9] Guo Huifeng, Tang Ruiming, Ye Yunming, *et al.* DeepFM: a factorization-machine based neural network for CTR prediction [C]// Proc of 26th International Joint Conference on Artificial Intelligence. 2017: 1725-1731.

[10] He Xiangnan, Liao Lizi, Zhang Hanwang, *et al.* Neural collaborative filtering [C]// Proc of International World Wide Web Conference Committee. 2017: 3-7.

[11] Luong M T, Pham H, Manning C D. Effective approaches to attention-based neural machine translation [C]// Empirical Methods on Natural Language Processing. 2015: 31-41.

[12] Mnih V, Heess N, Graves A. Recurrent models of visual attention [C]// Advances in Neural Information Processing Systems. 2014: 2204-2212.

[13] Zhang Qi, Gong Yeyun, Wu Jindou, *et al.* Retweet prediction with attention-based deep neural network [C]// Proc of the 25th ACM International on Conference on Information and Knowledge Management. New York: ACM Press, 2016: 75-84.

[14] Harper F M, Konstan J A. The MovieLens datasets: history and context [J]. ACM Trans on Interactive Intelligent Systems, 2016, 5 (4): 19.

[15] Hu Yifan, Koren Y, Volinsky C. Collaborative filtering for implicit feedback datasets [C]// Proc of the 8th IEEE International Conference on Data Mining. Washington DC: IEEE Computer Society, 2008: 263-272.

[16] Pan Rong, Zhou Yunhong, Cao Bin, *et al.* One-class collaborative filtering [C]// Proc of the 8th IEEE International Conference on Data Mining. Washington DC: IEEE Computer Society, 2008: 502-511.

[17] Bayer I, He Xiangnan, Kanagal B, *et al.* A generic coordinate descent framework for learning from implicit feedback [C]// Proc of International Conference on World Wide Web. International World Wide Web Conferences Steering Committee, 2017: 1341-1350.

[18] He Xiangnan, Zhang Hanwang, Kan M Y, *et al.* Fast matrix factorization for online recommendation with implicit feedback [C]// Proc of International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM Press, 2016: 549-558.

[19] Rendle S, Freudenthaler C, Gantner Z, *et al.* BPR: Bayesian personalized ranking from implicit feedback [C]// Proc of the 25th Conference on Uncertainty in Artificial Intelligence. 2012: 452-461.

[20] Wang Hao, Wang Naiyan, Yeung D Y. Collaborative deep learning for recommender systems [C]// Proc of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press, 2015: 1235-1244.

[21] Forbes P, Zhu Mu. Content-boosted matrix factorization for recommender systems: experiments with recipe recommendation [C]// Proc of ACM Conference on Recommender Systems. New York: ACM Press, 2011: 261-264.

[22] Sarwar B, Karypis G, Konstan J, *et al.* Item-based collaborative filtering recommendation algorithms [C]// Proc of International Conference on World Wide Web. New York: ACM Press, 2001: 285-295.

[23] 刘海龙，李宝安，吕学强，等. 基于深度卷积神经网络的图像检索算法研究 [J]. 计算机应用研究, 2017, 34 (12): 3816-3819. (Liu Hailong, Li Baoan, Lu Xueqiang, *et al.* Image retrieval algorithm based on deep convolutional neural network [J]. Research of Computer Applications, 2017, 34 (12): 3816-3819.